



## Invited Paper

# Classification of fragments of objects by the Fourier masks pattern recognition system



Carolina Barajas-García<sup>a</sup>, Selene Solorza-Calderón<sup>a,\*</sup>, Josué Álvarez-Borrego<sup>b</sup>

<sup>a</sup> Matemáticas Aplicadas, Facultad de Ciencias, Universidad Autónoma de Baja California, Km. 103 Carretera Tijuana-Ensenada, Ensenada, B.C., C.P. 22860, Mexico

<sup>b</sup> División de Física Aplicada, Departamento de Óptica, Centro de Investigación Científica y de Educación Superior de Ensenada, Carretera Ensenada-Tijuana No. 3918, Fraccionamiento Zona Playitas, Ensenada, B.C., C.P. 22860, Mexico

## ARTICLE INFO

## Article history:

Received 3 October 2015

Received in revised form

20 January 2016

Accepted 21 January 2016

Available online 29 January 2016

## Keywords:

Pattern recognition

Image processing

Pattern recognition systems

Binary rings masks

Fragments of objects

## ABSTRACT

The automation process of the pattern recognition for fragments of objects is a challenge to humanity. For humans it is relatively easy to classify the fragment of some object even if it is isolated and perhaps this identification could be more complicated if it is partially overlapped by other object. However, the emulation of the functions of the human eye and brain by a computer is not a trivial issue. This paper presents a pattern recognition digital system based on Fourier binary rings masks in order to classify fragments of objects. The system is invariant to position, scale and rotation, and it is robust in the classification of images that have noise. Moreover, it classifies images that present an occlusion or elimination of approximately 50% of the area of the object.

© 2016 Elsevier B.V. All rights reserved.

## 1. Introduction

Reproducing the pattern recognition human functions is a great challenge and a very difficult task. The research community has been invested a lot of effort to create robots and automation systems to this purpose. The introduction of the classical matched filter (CMF) by Vander Lugt [1] in 1964 generated great interest and progress in the pattern recognition systems by joint transform correlators. Unfortunately, these filters are specialized to solve specific problems, for example one filter could have an excellent performance in the discrimination step and the signal-to-noise ratio but a low efficiency under non-homogeneous illumination [2–4]. Although, composite filters have been used, the problem of rotation, scale and translation (RST) invariant correlator image descriptor is an active field of study due to its intrinsic complexity [5–11].

The scale invariant feature transform (SIFT) [5,6] and their variants [7,12–14] are robust and efficient local invariant feature descriptors for gray-level images. Local feature descriptors are used in a variety of pattern recognition real-world applications due to the identification efficiency of objects with moderate geometric distortions or partial occlusions. However, the performance of the local feature descriptors decays drastically when images

have noise or non-homogeneous illumination [9,10].

Recently, pattern recognition systems based on binary rings masks were developed [9,11,15,16]. These methodologies are robust and efficient in the gray-level images pattern recognition regardless of the position, rotation and, in some cases, objects' scale. Also, the response of these systems is great under non-homogeneous illumination and noise. In Ref. [15] the invariance to scale is achieved via the 2D non-separable scale transform. This 2D transform is not invariant to translation, hence the center of mass of the object is used to solve it. Based on the modulus of 2D transform of the image a single binary rings mask is built, therefore 1D RST signature is obtained. To avoid the problem of calculating the center of mass of the object, this work is based on the pattern recognition systems developed in Refs. [9,11,16], the systems utilized the amplitude spectrum of the image to obtain the invariance to position. Also, in those works are set four approaches to build the masks yielding four 1D RT signatures for a given image. Because these systems are invariant to position and rotation only, in the present work the invariance to scale was incorporated by using the analytical Fourier–Mellin transform. Moreover, the systems in [9,11,15,16] do not work with images of fragments of objects, the pattern recognition system proposed classifies that type of images too.

This work presents a RST invariant pattern recognition system based on the Fourier binary rings masks methodology [16]. Using the amplitude spectrum of the Fourier transform in order to obtain the invariance to translation and the normalized analytic Fourier–

\* Corresponding author.

E-mail address: [selene.solorza@gmail.com](mailto:selene.solorza@gmail.com) (S. Solorza-Calderón).

Mellin transform (AFMT) [17] to achieve the invariance to scale. At difference of the pattern recognition systems developed in Refs. [9,11,15,16], the RST invariant pattern recognition system described in this work classifies images that present an occlusion or elimination of a portion of the object, moreover the Z-Fisher transform was used to develop a 95% confidence interval for each given image, allowing the development of a MatLab GUI (Graphical User Interface) for the RST invariant digital image pattern recognition classifier.

The rest of this work is organized as follows: Section 2 describes the procedure to build the binary rings masks. Section 3 exposes the methodology to obtain the signature invariant to rotation, scale and translation based on Fourier transform, the analytic Fourier–Mellin transform and binary rings masks. Section 4 explains the image acquisition samples used in this work. Section 5 presents the manner to obtain the classifier output planes with a confidence level at least of 95% using the Z-Fisher transform and also the RST invariant digital image pattern recognition classifier MatLab GUI. Section 6 presents the analysis of the pattern recognition system efficiency when images have noise. Section 7 exposes a comparison analysis of the pattern recognition system proposed and other pattern recognition systems. Finally, conclusions are given in Section 8.

## 2. The Fourier masks

The mask of a selected gray-level image  $I(x, y)$ ,  $x = 1, \dots, N$ ,  $y = 1, \dots, M$  can be built by taking the real and imaginary parts of its Fourier transform [16], that is,  $Re(FT(I(x, y)))$  and  $Im(FT(I(x, y)))$ ; for example, the real and imaginary parts of the Fourier transform of Fig. 1a are shown in Figs. 1b and c, respectively. Next, the image of  $Re(FT(I(x, y)))$  and  $Im(FT(I(x, y)))$  are filtered by the binary disk mask  $D(x, y)$ , defined like,

$$D(x, y) = \begin{cases} 1, & \text{if } d((c_x, c_y), (x, y)) \leq n, \\ 0, & \text{otherwise,} \end{cases} \quad (1)$$

where  $(c_x, c_y)$  is the center-pixel of the image,  $n = \min\{c_x, c_y\}$  and  $d(p, q)$  is the Euclidean-distance between  $p$  and  $q$  points, thus the  $D(x, y)$  image is centered in the  $(c_x, c_y)$ -pixel. Fig. 1d presents an example of that binary filter  $D(x, y)$  and the results of the filter process are shown in Figs. 1e and f. Mathematically these operations are given by

$$f_R(x, y) = D(x, y) \cdot Re(FT(I(x, y))), \quad (2)$$

$$f_I(x, y) = D(x, y) \cdot Im(FT(I(x, y))). \quad (3)$$

For the images  $f_R(x, y)$  and  $f_I(x, y)$ , 180 profiles of  $2n$ -pixels length that passes for  $(c_x, c_y)$  were obtained. They are separated by  $\Delta\theta = 1^\circ$ , sampling in this manner the entire disk. Figs. 1e and f show (in black solid line) the profile we have named the zero-degree profiles and denoted by  $P_R^0(x)$  and  $P_I^0(x)$ , respectively. In general, the profile equations are expressed like

$$P_R^\theta = f_R(x, y(x)), \quad (4)$$

$$P_I^\theta = f_I(x, y(x)), \quad (5)$$

where  $x = 1, \dots, n$ ,  $y(x) = m(x - x_1) + y_1$ ,  $m = (y_2 - y_1)/(x_2 - x_1)$  is the slope of  $y$ ,  $(x_1, y_1) = (c_x + r \cos \theta, c_y - r \sin \theta)$  and  $(x_2, y_2) = (c_x + r \cos(\theta + \pi), c_y + r \sin(\theta + \pi))$  are the two distinct end points of that line segment,  $r = \min\{c_x, c_y\}$  and  $\theta$  is the angle that  $y$  has according to the horizontal axis in the Cartesian plane (considering that the origin  $(0, 0)$  of the Cartesian plane is set at the

center pixel of the image  $(c_x, c_y)$ ). Next, the scalars  $s_R^\theta$  and  $s_I^\theta$  that represent the addition of the square of the intensity values in each profile are computed, that is,

$$s_R^\theta = \sum_{x=1}^n (P_R^\theta(x))^2, \quad (6)$$

$$s_I^\theta = \sum_{x=1}^n (P_I^\theta(x))^2, \quad (7)$$

and the profile whose sum has the maximum value will be selected, that is

$$\alpha_\beta = \max_{0 \leq \theta \leq 179} \{S_R^\theta\}, \quad T_R(x) = P_R^\beta(x), \quad (8)$$

$$\alpha_\gamma = \max_{0 \leq \theta \leq 179} \{S_I^\theta\}, \quad T_I(x) = P_I^\gamma(x), \quad (9)$$

where  $\beta$  and  $\gamma$  are the angle of the profile in  $f_R(x, y)$  and  $f_I(x, y)$  whose sum has the maximum value, respectively. Hence, those profiles are called the maximum energy profiles. For example, in Figs. 1e and f the maximum energy profiles are shown (in black-dashed line) for the real and imaginary parts of the Fourier transform of the image Fig. 1a. Also, those profiles are given in the Cartesian plane in Figs. 1g and h. These figures show the symmetry of  $T_R(x)$  and the antisymmetry of  $T_I(x)$  in the vertical axis  $x = c_x$ . Next, based on the maximum energy profile obtained by Eq. (8), two binary functions  $Z_{RP}(x)$  and  $Z_{RN}(x)$  are built by

$$Z_{RP}(x) = \begin{cases} 1, & \text{if } T_R(x) > 0, \\ 0, & \text{if } T_R(x) \leq 0, \end{cases} \quad (10)$$

$$Z_{RN}(x) = \begin{cases} 0, & \text{if } T_R(x) > 0, \\ 1, & \text{if } T_R(x) \leq 0, \end{cases} \quad (11)$$

where  $x = 1, \dots, n$ . Analogously, based on the maximum energy profile obtained by Eq. (9), the  $Z_{IP}(x)$  and  $Z_{IN}(x)$  binary functions are constructed like,

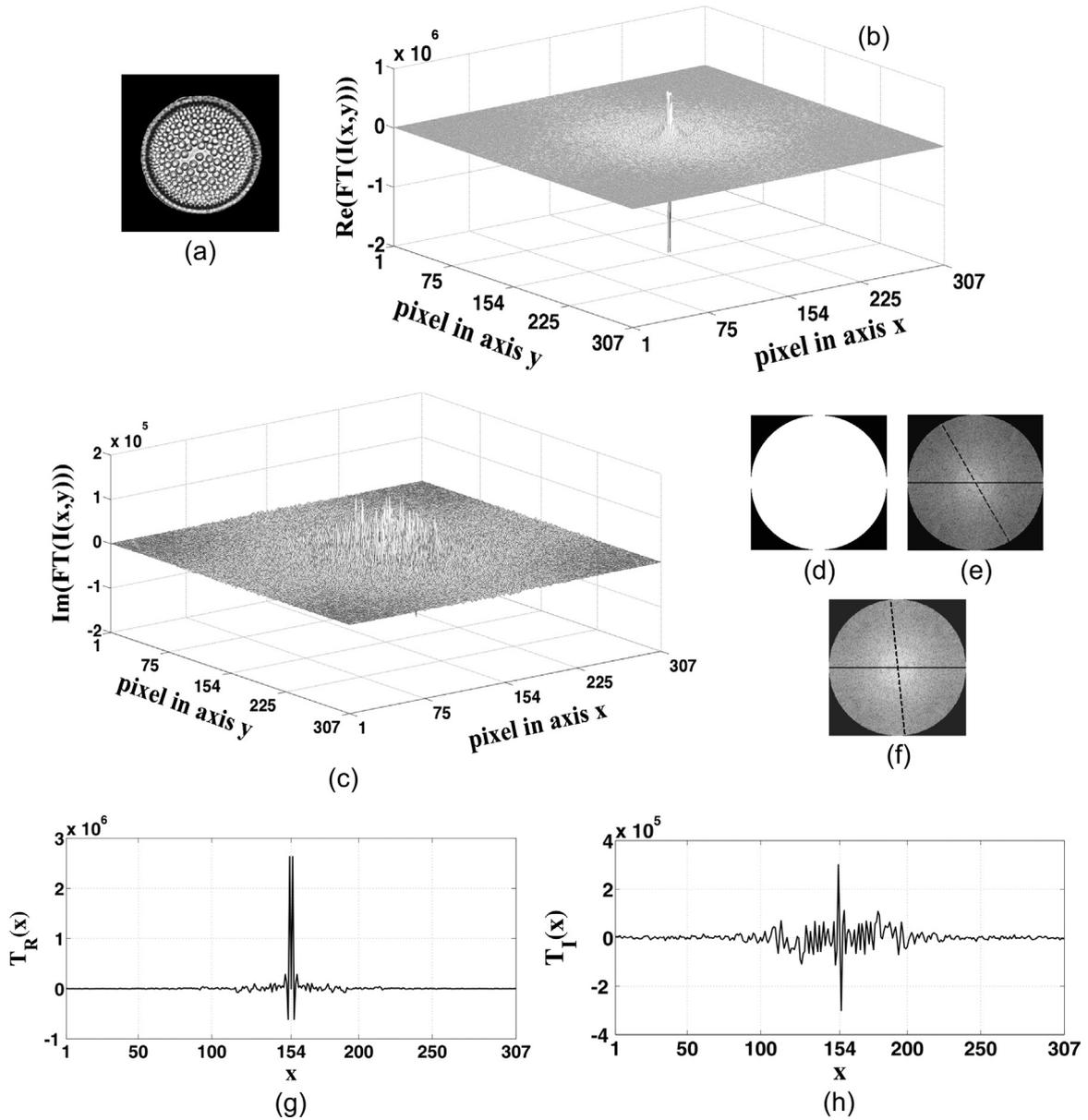
$$Z_{IP}(x) = \begin{cases} 1, & \text{if } T_I(x) > 0, \\ 0, & \text{if } T_I(x) \leq 0, \end{cases} \quad (12)$$

$$Z_{IN}(x) = \begin{cases} 0, & \text{if } T_I(x) > 0, \\ 1, & \text{if } T_I(x) \leq 0, \end{cases} \quad (13)$$

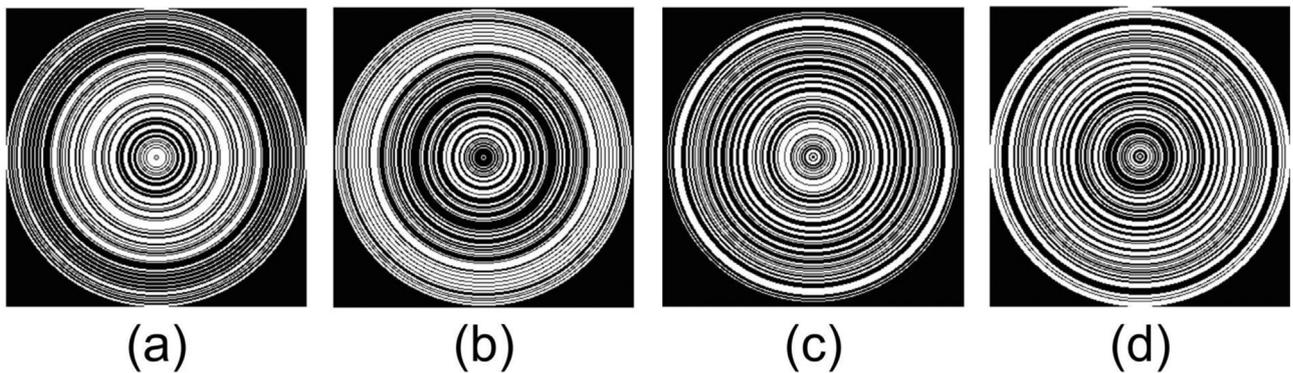
the first sub-index in Eqs. (10)–(13) indicate if the profile comes from the real (R) or the imaginary (I) part of the Fourier transform of the image. The second sub-index means that the positives values (P) or non-positives values (N) of the profile were taken. Finally, taking the vertical axis  $x = c_x$  as the rotation axis, the  $Z_{RP}(x)$ ,  $Z_{RN}(x)$ ,  $Z_{IP}(x)$  and  $Z_{IN}(x)$  functions are rotated  $360^\circ$  to obtain concentric cylinders of height one, different widths and centered in  $(c_x, c_y)$  pixel. Taking a cross-section of those concentric cylinders, the binary rings masks associated to the given image are built. Following the sub-index notation introduced for Eqs. (10)–(13), the binary rings masks are named  $M_{RP}(x, y)$ ,  $M_{RN}(x, y)$ ,  $M_{IP}(x, y)$  and  $M_{IN}(x, y)$ . Fig. 2 shows the binary rings masks corresponding to the image in Fig. 1a.

## 3. The signature

The pattern recognition system uses the amplitude spectrum  $A(u, v)$  of the Fourier transform of the image, because it is invariant



**Fig. 1.** (a) Image  $I(x, y)$ . (b) Real part of the Fourier transform of  $I(x, y)$ , that is  $Re(FT(I(x, y)))$ . (c) Imaginary part of the Fourier transform of  $I(x, y)$ , that is  $Im(FT(I(x, y)))$ . (d) Binary disk  $D(x, y)$ . (e)  $f_R(x, y) = D(x, y) \cdot Re(FT(I(x, y)))$ . The solid line shows the profile  $P_R^0(x)$  and the dashed line the profile  $T_R(x)$ . (f)  $f_I(x, y) = D(x, y) \cdot Im(FT(I(x, y)))$ . The solid line shows the profile  $P_I^0(x)$  and the dashed line the profile  $T_I(x)$ . (g) The maximum energy profile  $T_R(x)$ . (h) The maximum energy profile  $T_I(x)$ .



**Fig. 2.** (a) Mask  $M_{RP}$ . (b) Mask  $M_{RN}$ . (c) Mask  $M_{IP}$ . (d) Mask  $M_{IN}$ .

to translation [18]. Figs. 3a and b present the image  $I(x, y)$  and the corresponding amplitude spectrum  $|FT(I(x, y))|$ , respectively. Fig. 3c shows a translated version of  $I(x, y)$ , named  $I_T(x, y)$ , and Fig. 3d

exhibits the amplitude spectrum  $|FT(I_T(x, y))|$ . Because of  $|FT(I(x, y))| = |FT(I_T(x, y))|$ , the system is invariant to translation in an easy manner. Mathematically, it is given by

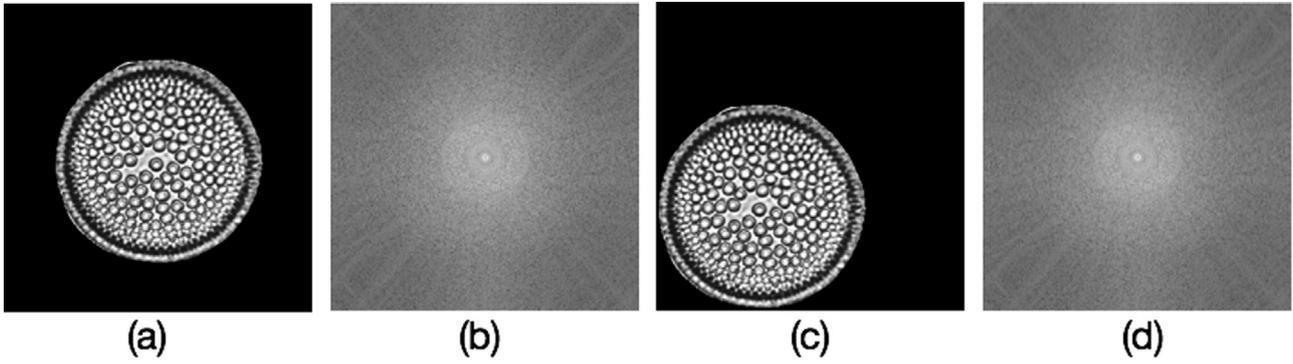


Fig. 3. (a)  $I(x, y)$ . (b)  $|FT(I(x, y))|$ . (c)  $I_T(x, y)$ . (d)  $|FT(I_T(x, y))|$ .

$$A(u, v) = |FT(I(x, y))| = \sqrt{\text{Re}^2(FT(I(x, y))) + \text{Im}^2(FT(I(x, y)))}. \quad (14)$$

$$M(k, \omega) = \mathcal{M}\{A(e^\rho, \theta)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_0^{2\pi} A(e^\rho, \theta) e^{\rho\sigma} e^{-i(k\theta + \rho\omega)} d\theta d\rho, \quad (15)$$

The next step is the invariance to scale, which is obtained via the fast analytical Fourier–Mellin transform (AFMT), given by

where  $\rho = \ln(r)$  and  $\sigma > 0$ . Fig. 4c shows  $A(u, v)$  in log-polar coordinates as it is required in Eq. (15), however this equation is not invariant to scale, but normalizing the AFMT by its value in the

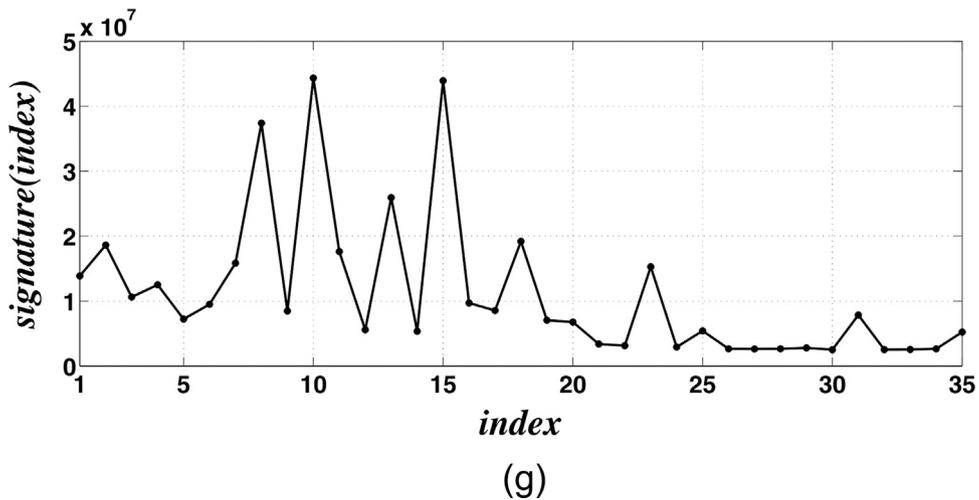
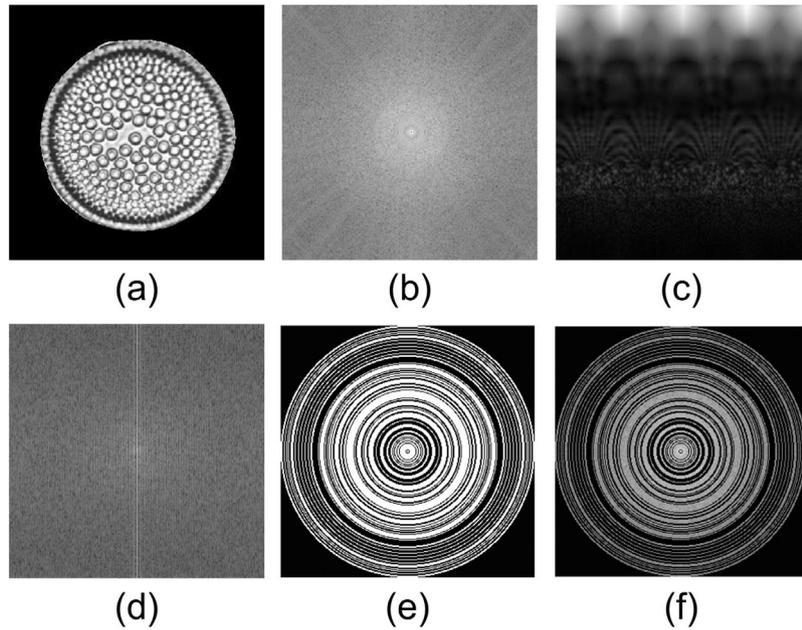


Fig. 4. (a) Image  $I(x, y)$ . (b)  $A(u, v)$ . (c)  $A(e^\rho, \theta)$ . (d)  $S(k, \omega)$ . (e) Binary mask  $M_{RP}$ . (f)  $H_{RP}(x, y)$ . (g) Signature of image  $I(x, y)$ .

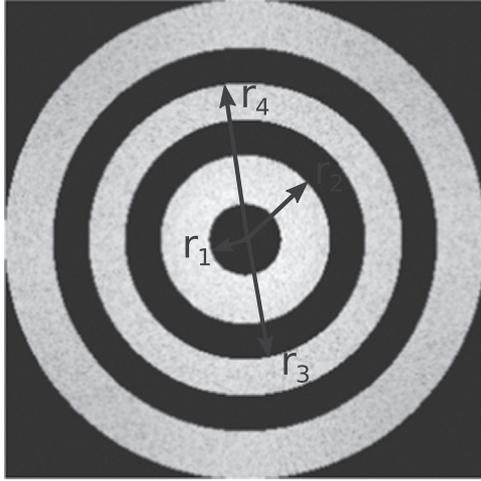


Fig. 5. Example of the numbering process of the rings in a binary mask.

center-pixel, the amplitude spectrum is invariant to scale [17], that is

$$S(k, \omega) = \left| \frac{M(k, \omega)}{M(c_x, c_y)} \right|, \quad (16)$$

where  $(c_x, c_y)$  is the central pixel of the image  $M(k, \omega)$ . The next step is to filter the  $S(k, \omega)$  amplitude spectrum (Fig. 4d) of the image (Fig. 4a) by a binary ring mask, for example  $M_{RP}(x, y)$  (Fig. 4e) as

$$H_{RP} = M_{RP} \circ S, \quad (17)$$

where  $\circ$  means an element-wise product or Hadamard product [19]. The results of Eq. (17) are presented in Fig. 4f.

The rings in  $H_{RP}(x, y)$  are enumerated from inside to outside to obtain the following set,

$$ring_k = \{(x, y) : x = c_x + r \cos \theta, y = c_y - r \sin \theta; r_{2k-1} \leq r \leq r_{2k}\}, \quad (18)$$

where  $r_i$  is the distance from the center to the  $i$ th ring. Fig. 5 shows how the rings are numbered to form the set in Eq. (18). The band lying between  $r_1$  and  $r_2$  is taken as the first ring, the second ring is set between  $r_3$  and  $r_4$ , and successively, therefore the band that lies between  $r_{2k-1}$  and  $r_{2k}$  is named the  $k$ th ring. After that, the addition of the intensity values in each ring of  $H_{RP}(x, y)$  are computed to build the function

$$signature(index) = \sum H_{RP}(x, y), \text{ if } H_{RP}(x, y) \in ring_{index}, \quad (19)$$

where  $index = 1, \dots, m$  and  $m$  represents the number of rings in  $H_{RP}$ . Fig. 4g shows the signature constructed by the binary rings mask  $M_{RP}$ , to follow the nomenclature introduced to call the masks, it is named  $S_{RP}$ . Analogously,  $S_{RN}$ ,  $S_{IP}$  and  $S_{IN}$  are obtained using  $M_{RN}$ ,  $M_{IP}$  and  $M_{IN}$ , respectively. Fig. 6 presents the four signatures of the image in Fig. 4a.

#### 4. Image acquisition

To test the system, this work uses a database of 18 gray-level digital images of fossil diatoms (Fig. 7). Those images were selected because of the similarity in their morphologies. The samples were collected in 1996 by an oceanographic survey in La Cuenca San Lázaro, Baja California Sur, México [20]. To generate a database of problem images of fragments of objects, each diatom in Fig. 7 was fragmented by hand to obtain 49 images of each diatom with

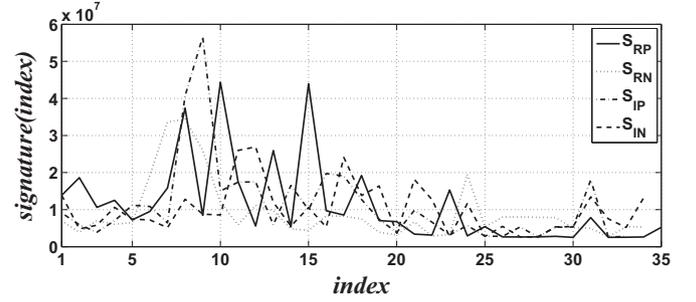


Fig. 6. The four signatures of Fig. 4a.

different percentages of missing data: from 1% to 99%. For example, Fig. 8 shows the gray-level digital images for *Actinocyclus ingens-Rattray* diatom.

#### 5. Classification

The pattern recognition system works comparing the signatures of the image to be classified (problem image) with the signatures of each reference images in the database  $\beta_R = \{R_j : j = 1, \dots, k; k \in \mathbb{N}\}$ . In this work  $\beta_R$  are the diatom images in Fig. 7 and the four signatures  $S_{R_j}^{RP}$ ,  $S_{R_j}^{RN}$ ,  $S_{R_j}^{IP}$  and  $S_{R_j}^{IN}$  are obtained to each  $R_j$  image. After that, the signature  $\hat{S}_{R_j}$  is calculated like

$$\hat{S}_{R_j} = S_{R_j}^{RP} \circ S_{R_j}^{RN} \circ S_{R_j}^{IP} \circ S_{R_j}^{IN}, \quad (20)$$

to be processed as

$$\hat{S}_{R_j} = Re(FT(\hat{S}_{R_j})). \quad (21)$$

The final step is to determine the feature that will characterize the pattern in the target image  $R_j$ . This is calculated by applying Pearson's correlation coefficient [21],

$$r_{R_j} = \frac{\max \{ |C_L(\hat{S}_{R_j})| \}}{(N-1)\sigma_{\hat{S}_{R_j}}^2}, \quad (22)$$

where  $N$  is the cardinality of the domain of  $\hat{S}_{R_j}$  and  $\sigma_{\hat{S}_{R_j}}$  is the standard deviation of the signature.  $C_L$  represents the linear correlation of two given signatures  $S_1$  and  $S_2$ , that is

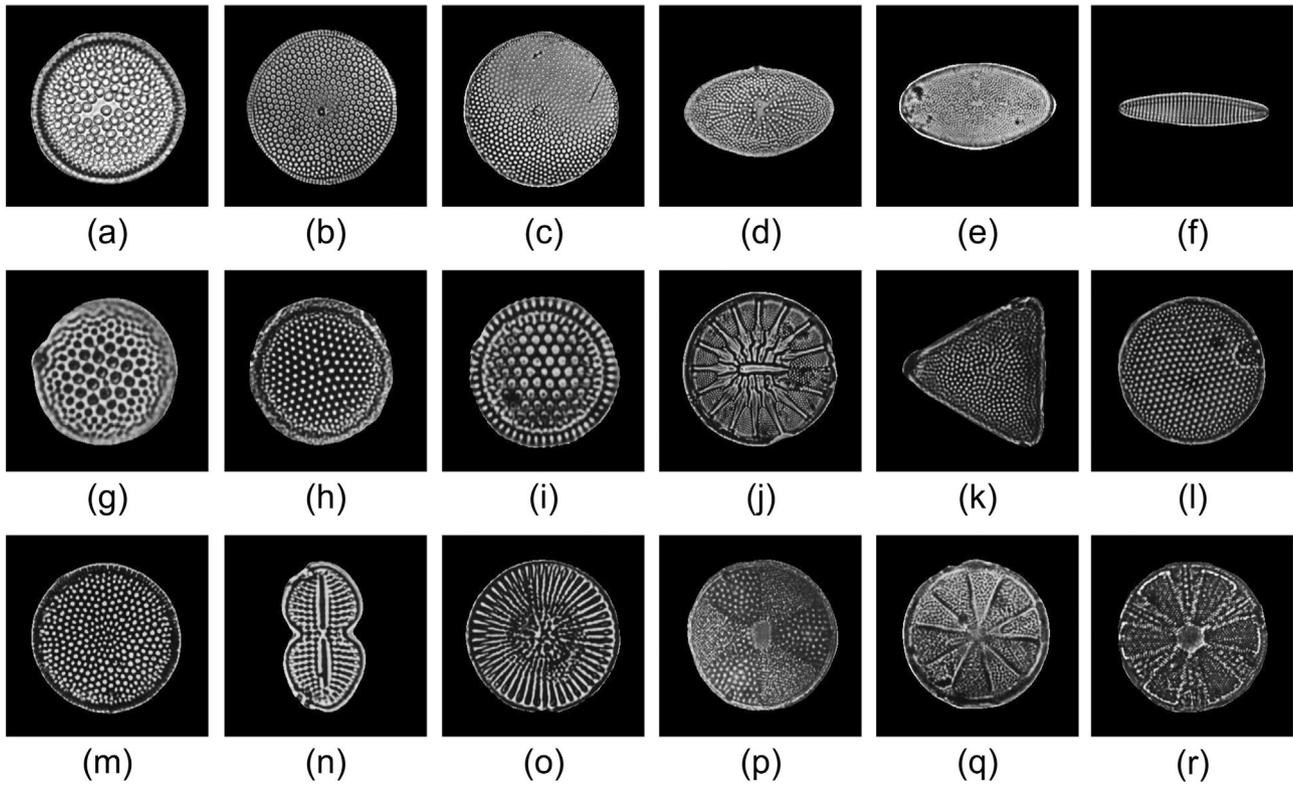
$$C_L(S_1, S_2) = FT^{-1} \{ |FT(S_2)| e^{i\phi} |FT(S_1)| e^{-i\varphi} \}, \quad (23)$$

where  $\varphi$  and  $\phi$  are the phases of the Fourier transform of the signatures  $S_1$  and  $S_2$ , respectively. The notation  $C_L(S_1)$  indicates the autocorrelation function. Analogously, the feature that will characterize the pattern in the problem image  $P$  is set by

$$r_P = \frac{\max \{ |C_L(\hat{S}_{R_j}, \hat{S}_P)| \}}{(N-1)\sigma_{\hat{S}_{R_j}} \sigma_{\hat{S}_P}}, \quad (24)$$

where  $\hat{S}_P$  is the signature of the problem image and  $\sigma_{\hat{S}_P}$  is the standard deviation of that signature. If  $r_P$  is similar to  $r_{R_j}$ , then  $P$  and  $R_j$  are the same, otherwise they are different.

In order that the pattern recognition digital system also classifies images with fragments of objects, the system was trained using 50 images with different percentages of missing data, for example the diatoms *Actinocyclus ingens-Rattray* in Fig. 8. Then, their corresponding Pearson's correlation coefficients  $r_{R_j}^k$ ,  $k = 0, \dots, 49$ , were



**Fig. 7.** Data base of 18 diatoms: (a) *Actinocyclus ingens* - Ratray. (b) *Azpeitia* sp. (c) *Azpeitia nodulifera* - (Schmidth) Fryxell et Sims. (d) *Actinocyclus ellipticus* - Grunow in van Heurck. (e) *Actinocyclus ellipticus* var *moronensis* - (Deby ex Ratray) Kolbe. (f) *Nitzschia praereinholdii* - Schrader. (g) *Thalassiosira oestruppii* var 1. (h) *Thalassiosira oestruppii* var 2. (i) *Thalassiosira domifacta* - (Hendey) Jouse. (j) *Asteromphalus imbricatus* - Wallich. (k) *Pseudotriceratium cinnamomeum* - (Greville) Grunow. (l) *Thalassiosira kozlovii* - Makarova. (m) *Coscinodiscus radiatus* - Ehrenberg. (n) *Diploneis bombus* - Cleve-Euler in Backman et Cleve-Euler. (o) *Stephanodiscus* sp. (p) *Actinoptochus undulatus* - (Bailey) Ralf. (q) *Actinoptochus bipunctatus* - Lohman. (r) *Actinoptochus splendens* - (Shadbolt) Ralf ex Pritchard.

obtained. Because  $\{r_{R_j}^k, k = 0, \dots, 49\}$  does not have a normal distribution, those values are normalized by the Z-Fisher transform to get the confidence interval for the correlation values [21]. The Z-Fisher value for  $r_{R_j}^k$  is given by

$$Z_{r_{R_j}^k} = 1.1513 \ln \left( \frac{1 + r_{R_j}^k}{1 - r_{R_j}^k} \right). \quad (25)$$

Thus, the 95% confidence interval for  $Z_{r_{R_j}^k}$  is

$$\left[ Z_{r_{R_j}^k}^-, Z_{r_{R_j}^k}^+ \right] = \left[ Z_{r_{R_j}^k} - 1.96\sigma_Z, Z_{r_{R_j}^k} + 1.96\sigma_Z \right], \quad (26)$$

with a standard deviation of  $\sigma_Z = 1/\sqrt{n-3}$  and  $n=50$  the size of the sample. Hence, the confidence interval for the correlation coefficient  $\rho_{r_{R_j}^k}$  is

$$\rho_{r_{R_j}^k}^- \leq \rho_{r_{R_j}^k} \leq \rho_{r_{R_j}^k}^+, \quad (27)$$

where

$$\rho_{r_{R_j}^k}^- = \frac{\exp\left(2Z_{r_{R_j}^k}^-\right) - 1}{\exp\left(2Z_{r_{R_j}^k}^-\right) + 1}, \quad \rho_{r_{R_j}^k}^+ = \frac{\exp\left(2Z_{r_{R_j}^k}^+\right) - 1}{\exp\left(2Z_{r_{R_j}^k}^+\right) + 1}. \quad (28)$$

For each  $R_j$  there are 50 values of  $\rho_{r_{R_j}^k}^-$  and another 50 values for  $\rho_{r_{R_j}^k}^+$  then the confidence interval of 95% to decide if a problem image and  $R_j$  are the same is given by

$$\left[ \min_{0 \leq k \leq 49} \left\{ \rho_{r_{R_j}^k}^- \right\}, \max_{0 \leq k \leq 49} \left\{ \rho_{r_{R_j}^k}^+ \right\} \right]. \quad (29)$$

The problem image database has 900 samples of real fragments of diatoms. Table 1 resumes the results obtained by the Fourier pattern recognition system, the column *Diat.* indicates the diatom fossil type, the column *M.P.* points the minimum percentage (M.P.) of the area of the object that is required to its classification with a confidence level at least of 95% and the column *Image* shows the corresponding image with the minimum percentage. Table 1 indicates that the system is robust and properly classifies images that present an occlusion or elimination until 49% of the area of the object. Moreover, the Z-Fisher methodology allows to assign a 95% confidence interval to each problem image, thus the automation of the classification step could be done. A MatLab GUI (Graphical User Interface) was developed for the Fourier pattern recognition system described in this work. In the GUI example of Fig. 9, the problem image (PI) is a fragment of the reference image (RI), however the signatures of both images are quite similar (left-bottom part of the GUI), hence the confidence interval of the PI is contained into the confidence interval of the reference image (right-bottom part of the GUI), indicating that the problem image is equal to the reference image.

## 6. Noise analysis

To test the performance of the Fourier pattern recognition system, the discrimination coefficient *DC* [4] was used and it is defined as

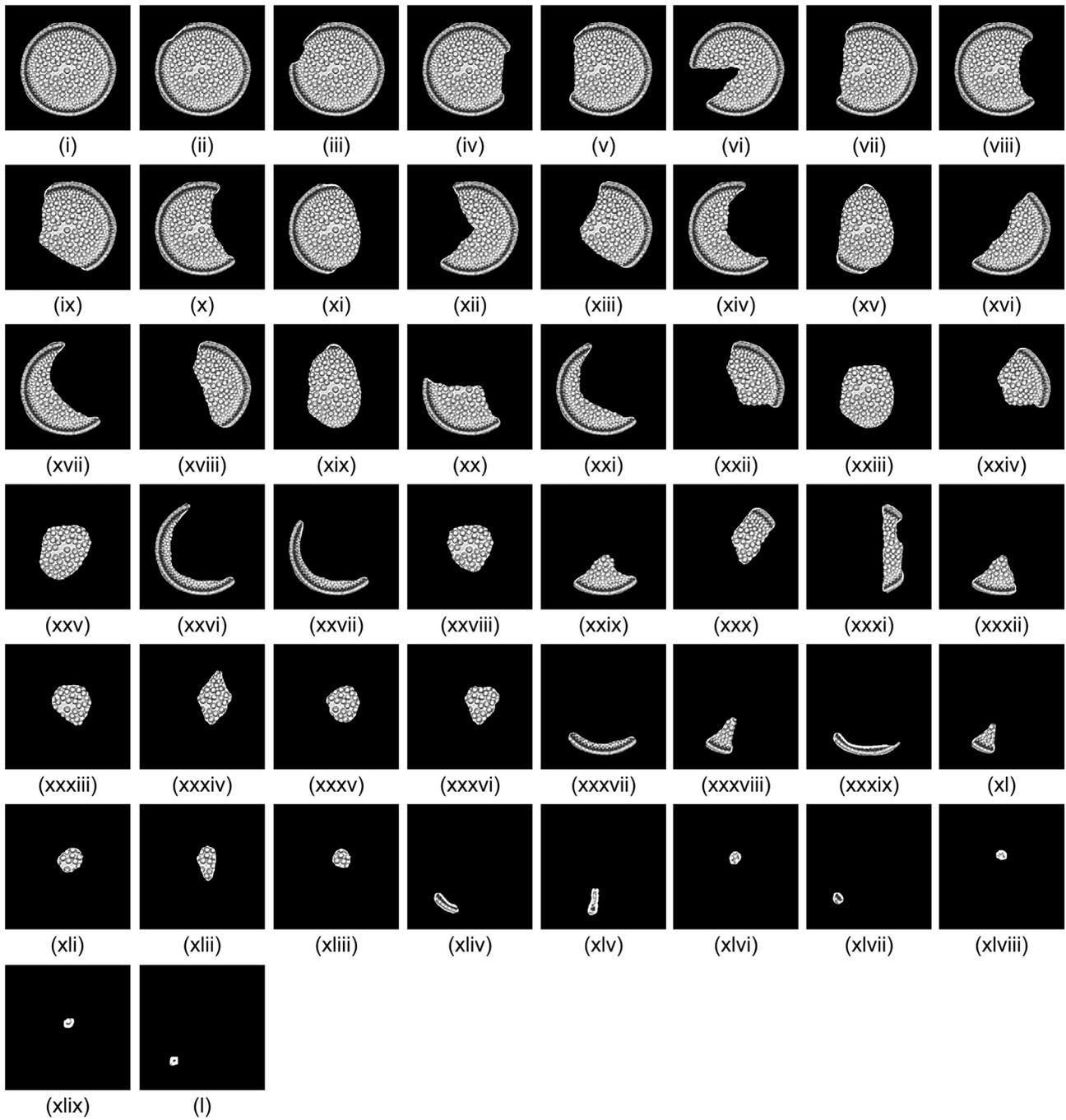


Fig. 8. Problem images of *Actinocyclus ingens-Ratray* diatom.

$$DC = 1 - \frac{\max |C_L(S_T, S_N)|^2}{(Q(0))^2}, \tag{30}$$

where  $Q = |C_L(S_T, S_{TN})|$  and  $S_T$ ,  $S_{TN}$  and  $S_N$  are the signatures of the target, the target with noise and the background image with noise, respectively. For the sake of comparison, the performance of the SURF methodology is included, but here the results are given in terms of the repeatability parameter  $r$ ,

$$r = \frac{C(T, P)}{\text{mean}(N_T, N_P)}, \tag{31}$$

where  $C(T, P)$  represents the number of the common detected points in the reference image  $T$  and the problem image  $P$ ;  $N_T$  and  $N_P$  are the number of points detected in  $T$  and  $P$ , respectively [22].

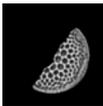
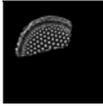
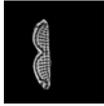
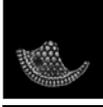
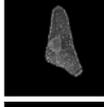
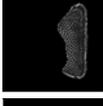
Fig. 10 presents the graphs of the mean of the  $DC$  response for

the Fourier pattern recognition system and the repeatability analysis ( $r$  values) for the SURF algorithm, the images were altered with additive Gaussian noise of media zero and variance from zero to 1, using 50 images per sample. In Fig. 10 is shown that the Fourier system has better response than the SURF methodology. The same analysis was done using salt and pepper noise, obtaining the same result as the additive Gaussian noise, Fig. 11.

### 7. Discussions

In this work a pattern recognition digital system that is invariant to translation, scale and rotation is presented. The system accepts an scale range until  $\pm 10\%$  with respect to the reference image. Also, the system showed to be robust to classify images

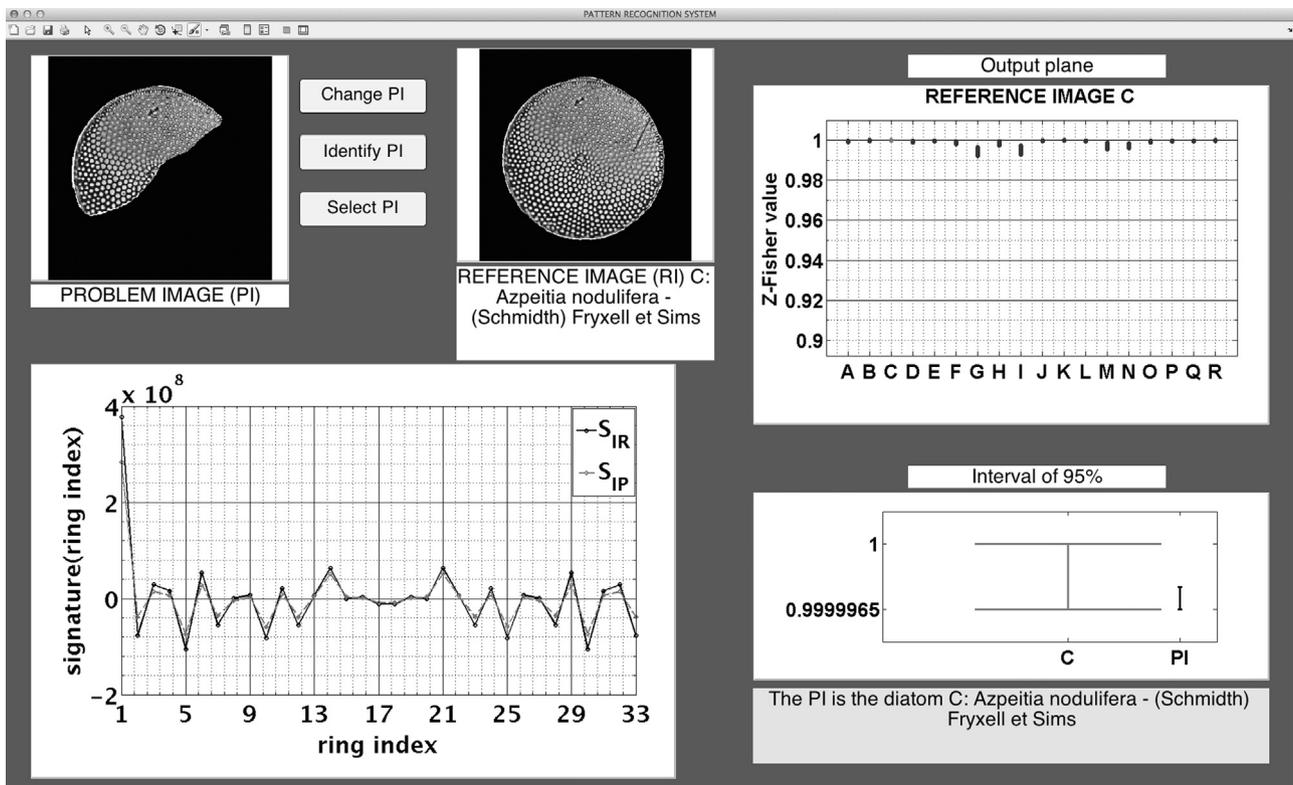
**Table 1**  
Minimum percentage fragment required.

Diat.	M.P.	Image	Diat.	M.P.	Image	Dia.	M.P.	Image
a	48%		g	48%		m	34%	
b	22%		h	39%		n	37%	
c	25%		i	39%		o	34%	
d	27%		j	39%		p	32%	
e	29%		k	47%		q	25%	
f	44%		l	39%		r	49%	

with Gaussian and salt and pepper noises. Moreover, it classifies images even if they contain fragments up to 49% of the area of the object.

Other technique based on 1D RST invariant signatures is presented in [10]. This methodology, called vectorial signatures, uses the non-linear correlation function (with  $k=0.3$ ) to compare the

signatures. The vectorial signatures system was tested with the reference image database in Fig. 7, showing an excellent performance to classify images until  $\pm 20\%$  in the scale range. For example, Fig. 12a presents one of the eighteen output planes generated by the system, there the reference image is diatom *Actinocyclus ingens-Rattray*. Fig. 12b exhibits an amplification zone



**Fig. 9.** MatLab GUI of the Fourier pattern recognition system.

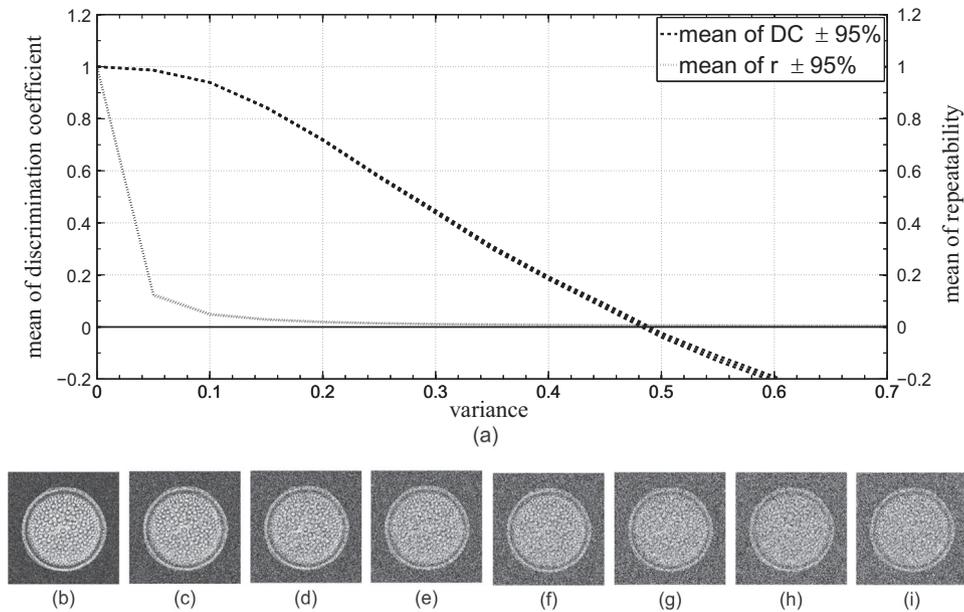


Fig. 10. Pattern recognition systems performance when images have additive Gaussian noise.

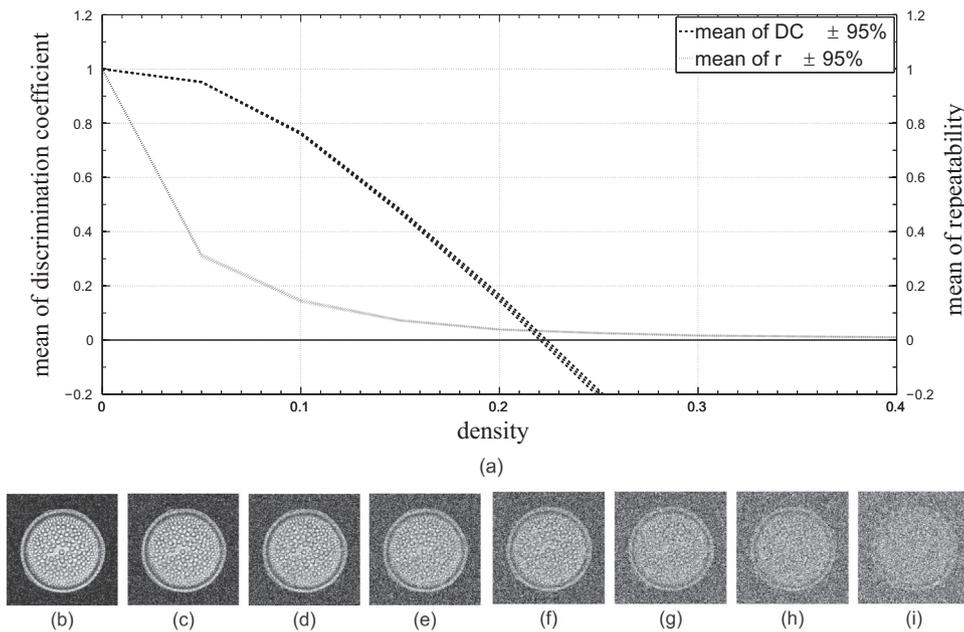
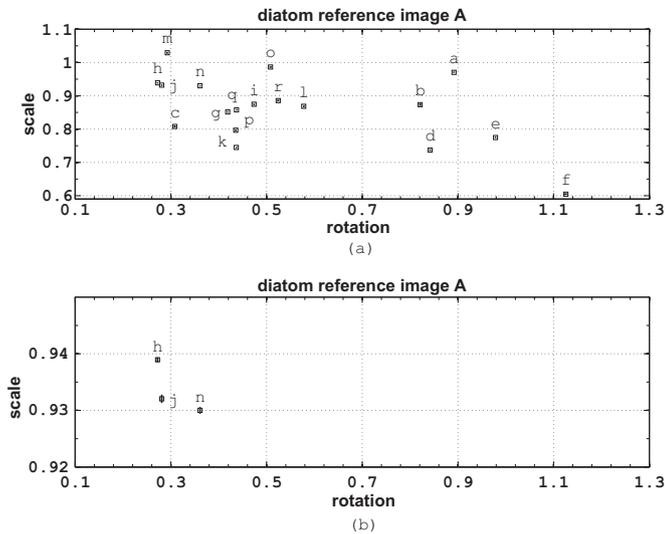


Fig. 11. Pattern recognition systems performance when images have salt and pepper noise.

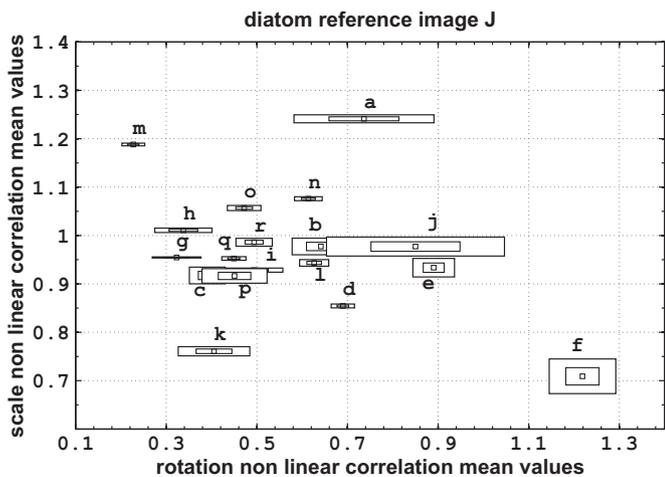
around h and j diatoms' boxes to show that none of the boxes are overlapped. The same results were obtained with the other seventeen reference images in the database, therefore the system has a confidence level at least of 95.4%. However, when this system was tested using the images with the fragments of diatoms (for example Fig. 8), it could not classify the diatoms that missed more than 10% of the diatom's area. Fig. 13 shows the output plane for the reference image j and problem images with diatoms without until 10% of the area. It is observed that the boxes for diatoms c, i and p are overlapped and the boxes corresponding to diatoms b and j overlap too. Hence, the vectorial signatures system does not work with images that have a fragment of the object.

On the other hand, the SURF methodology does not work properly when the image was rotated. Fig. 14 shows the repeatability responses of this system (the parameter  $r$  given in Eq. (31)). The diatom image named a in Fig. 7 and their 360 rotated version (the image was rotated degree by degree until complete the circle)

were used for this test. Fig. 14 exhibits that at a very low rotation angle, the system efficiency decays notoriously because of the presence of the *sawtooth* noise that affects the system performance drastically. Also, it shows a periodic behavior of  $90^\circ$ ; when the image was rotated  $45^\circ$  the system has the lowest performance since the *sawtooth* noise is greater. On the contrary, at rotation angles near  $90^\circ$  the parameter  $r$  tends to 1, because the images have lesser noise effect. The pattern recognition system invariant to position, scale and rotation proposed in this work classifies that kind of noisy images with a confidence level at least of 95%, therefore it is showed to be robust in the classification of images with *sawtooth* noise also. Although the system is specialized to work with gray-level digital images that have just one object, this is not a limitation because if the image has more than one object, a preprocessing technique could be used to split it in several images that contains just one object per image. Currently, the authors are working to extend the applicability of this system to color images.



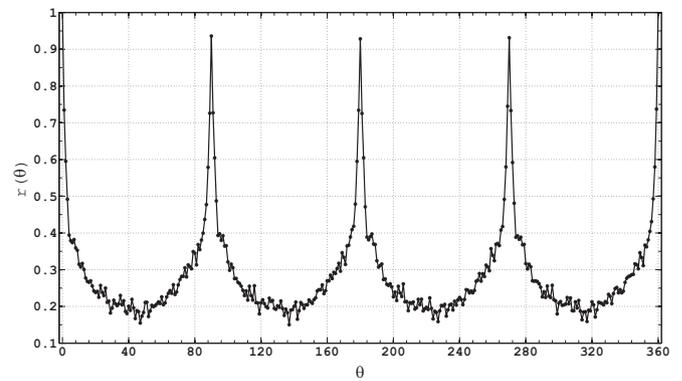
**Fig. 12.** The output plane obtained using the pattern recognition system in [10]. The horizontal axis shows the mean of the maximums of the non-linear correlation of the rotation signatures  $\pm 2$  standard errors. The vertical axis shows the mean of the maximums of the non-linear correlation of the scale signatures  $\pm 2$  standard errors.



**Fig. 13.** The output plane obtained using the pattern recognition system in [10]. The horizontal axis shows the mean of the maximums of the non-linear correlation of the rotation signatures  $\pm 2$  standard errors. The vertical axis shows the mean of the maximums of the non-linear correlation of the scale signatures  $\pm 2$  standard errors.

## 8. Conclusions

The pattern recognition system invariant to translation, scale and rotation proposed in this work shows an excellent performance, a confidence level at least of 95%, in the classification of gray-level images even if they contain fragments up to 49% of the area of the object. This pattern recognition system is based on the Fourier binary rings mask methodology invariant to translation and rotation, which is robust and efficient under Gaussian, salt and pepper, *sawtooth* noises, non-homogeneous illumination. In this work, the invariance to scale was introduced in the system via the analytical Fourier–Mellin transform, obtaining a scale range of  $\pm 10\%$ . Also, the use of the Z-Fisher statistical methodology allows us to assign a 95% confidence level to each image, yielding the automation of the classification step. Therefore, a MatLab GUI was developed to automate the classification of digital images with a fragment of the object. The methodology proposed in this work



**Fig. 14.** The repeatability response of SURF methodology, Eq. (31).

was compared with SURF technique and the 1D vectorial signatures methodology in [10]. The vectorial signatures system does not classify images with fragments of objects and the SURF system does not work properly with images that have few *sawtooth* noise. Although, the *sawtooth* noise were introduced by rotating the images, this kind of noise is presented in low resolution images too. Hence, the pattern recognition system based on binary rings masks is an excellent option to be used in the classification of gray-level digital images with one object. If the image has more than one object, a segmentation technique should be used to split the image in several images containing one object.

## Acknowledgments

This work was partially supported by CONACyT under Grant no. 169174. Carolina Barajas-García was a student in the M.C. program MyDCI offered by Universidad Autónoma de Baja California and she was supported by CONACyT scholarship. The grammatical critical reviews of the manuscript by LD.I. Horacio Padilla Calderón and Ing. Virgilio E. Padilla Calderón are greatly appreciated. We also thank the unknown reviewers for the excellent comments to improve this manuscript.

## References

- [1] A.V. Lugt, Signal detection by complex spatial filtering, *IEEE Trans. Inf. Theory* 10 (2) (1964) 139–145, <http://dx.doi.org/10.1109/TIT.1964.1053650>.
- [2] J.L. Horner, P.D. Gianino, Phase-only matched filtering, *Appl. Opt.* 23 (6) (1984) 812–816, <http://dx.doi.org/10.1364/AO.23.000812>.
- [3] C.F. Hester, D. Casasent, Multivariate technique for multiclass pattern recognition, *Appl. Opt.* 19 (11) (1980) 1758–1761, <http://dx.doi.org/10.1364/AO.19.001758>.
- [4] B. Vijaya Kumar, L. Hassebrook, Performance measures for correlation filters, *Appl. Opt.* 29 (20) (1990) 2997–3006, <http://dx.doi.org/10.1364/AO.29.002997>.
- [5] D.G. Lowe, Object recognition from local scale-invariant features, in: *The Proceedings of the Seventh IEEE International Conference on Computer Vision*, vol. 2, IEEE, Kerkyra, Greece, 1999, pp. 1150–1157, <http://dx.doi.org/10.1109/ICCV.1999.790410>.
- [6] D.G. Lowe, Distinctive image features from scale-invariant keypoints, *Int. J. Comput. Vis.* 60 (2) (2004) 91–110, <http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94>.
- [7] H. Bay, T. Tuytelaars, L. Van Gool, SURF: Speeded Up Robust Features, in: *Computer Vision—ECCV*, Springer, Germany, 2006, pp. 404–417, [http://dx.doi.org/10.1007/11744023\\_32](http://dx.doi.org/10.1007/11744023_32).
- [8] J.R. Lerma Aragón, J. Álvarez-Borrego, Vectorial signatures for invariant recognition of position, rotation and scale pattern recognition, *J. Mod. Opt.* 56 (14) (2009) 1598–1606, <http://dx.doi.org/10.1080/09500340903203111>.
- [9] S. Solorza, J. Álvarez-Borrego, Digital system of invariant correlation to position and rotation, *Opt. Commun.* 283 (19) (2010) 3613–3630, <http://dx.doi.org/10.1016/j.optcom.2010.05.035>.
- [10] C. Fimbres-Castro, J. Álvarez-Borrego, M.A. Bueno-Ibarra, Invariant nonlinear correlation and spectral index for diatoms recognition, *Opt. Eng.* 51 (4) (2012) 047201-1, <http://dx.doi.org/10.1117/1.OE.51.4.047201>.
- [11] J. Álvarez-Borrego, S. Solorza, M.A. Bueno-Ibarra, Invariant correlation to position and rotation using a binary mask applied to binary and gray images, *Opt. Commun.* 294 (2013) 105–117, <http://dx.doi.org/10.1016/j.optcom.2012.12.010>.

- [12] Y. Ke, R. Sukthankar, PCA-SIFT: A more distinctive representation for local image descriptors, in: *The Proceedings of the Conference on Computer Vision and Pattern Recognition*, vol. 2, IEEE, Washington, DC, USA, 2004, pp. II–506.
- [13] E.N. Mortensen, H. Deng, L. Shapiro, A SIFT descriptor with global context, in: *The Proceedings of the Conference on Computer Vision and Pattern Recognition*, vol. 1, IEEE, San Diego, CA, USA, 2005, pp. 184–190. <http://dx.doi.org/10.1109/CVPR.2005.45>.
- [14] D. Su, J. Wu, Z. Cui, V.S. Sheng, S. Gong, CGCI-SIFT: a more efficient and compact representation of local descriptor, *Meas. Sci. Rev.* 13 (3) (2013) 132–141, <http://dx.doi.org/10.2478/msr-2013-0022>.
- [15] A.S. Ventura, J.Á. Borrego, S. Solorza, Adaptive nonlinear correlation with a binary mask invariant to rotation and scale, *Opt. Commun.* 339 (2015) 185–193, <http://dx.doi.org/10.1016/j.optcom.2014.11.051>.
- [16] S. Solorza, J. Álvarez-Borrego, Position and rotation-invariant pattern recognition system by binary rings masks, *J. Mod. Opt.* 62 (10) (2015) 851–864, <http://dx.doi.org/10.1080/09500340.2015.1013579>.
- [17] S. Derrode, F. Ghorbel, Robust and efficient Fourier–Mellin transform approximations for gray-level image reconstruction and complete invariant description, *Comput. Vis. Image Underst.* 83 (1) (2001) 57–78, <http://dx.doi.org/10.1006/cviu.2001.0922>.
- [18] H.P. Hsu, R.G.F. Torrez, *Análisis de Fourier*, Addison-Wesley Iberoamericana, Edo. de Mexico, Mexico, 1987.
- [19] G.H. Golub, C.F. Van Loan, *Matrix Computations*, vol. 3, JHU Press, Maryland, USA, 2012.
- [20] M. Esparza-Álvarez, Variabilidad de la comunidad de diatomeas en los sedimentos de La Cuenca de San Lázaro, Baja California Sur; México, Tesis de Maestría. CICESE, Ensenada, B.C., México.
- [21] A. Sánchez-Bruno, A. Borges del Real, Transformación Z de Fisher para la determinación de intervalos de confianza del coeficiente de correlación de Pearson, *Psicothema* 17 (1) (2005) 148–153.
- [22] L. Juan, O. Gwun, A comparison of SIFT, PCA-SIFT and SURF, *Int. J. Image Process.* 3 (4) (2009) 143–152.